2014

Linguistic Capital, Information Access and Economic Opportunity among Rural Young Adults in Western China*

Emily C. Hannum
University of Pennsylvania, hannumem@soc.upenn.edu

Huayu Sebastian Cherng
New York University, cherng@nyu.edu

Linguistic Capital, Information Access and Economic Opportunity among Rural Young Adults in Western China*

Abstract
Facility with a country's dominant language, a key form of linguistic capital, has a role to play in processes of social stratification and mobility, and this role is poorly understood. We have sought, in this paper, to explore access to this form of linguistic capital, and the implications of possessing linguistic capital, for a group of young adults who have been historically disadvantaged: rural young adults in western minority areas. Three main results emerge. First, there is a great deal of variability in linguistic capital, defined as standard Mandarin facility, across provinces and ethnic groups covered in the CHES sample. The greatest gap appears in Xinjiang, where Han residents have very high facility in standard Mandarin, and where minority residents report very low facility. In some provinces in the CHES sample, there are minimal differences between majority and minority populations. Second, standard Mandarin facility is tied to information access, in the form of Internet use. Facility in minority languages is not. Third, Mandarin facility, but not minority language facility, is linked to economic opportunity in young adulthood.

Keywords
language, information, inequality, rural, youth, China

Disciplines
Asian Studies | Bilingual, Multilingual, and Multicultural Education | Curriculum and Instruction | Demography, Population, and Ecology | Educational Assessment, Evaluation, and Research | Educational Sociology | First and Second Language Acquisition | International and Comparative Education | Linguistics | Race and Ethnicity

Comments
This version is a 2014 working paper. Manuscript subsequently published in:

Gustafsson, B., Hasmath, R. and Sai, D., eds. (2017) Ethnicity and Inequality in China. New York and Oxford: Routledge.

Chinese Version: 中国少数民族的差异性研究 (社会科学文献出版社)

Linguistic Capital, Information Access and Economic Opportunity among Rural Young Adults in Western China*

May 29, 2014

Emily Hannum*
hannumem@soc.upenn.edu

Hua-Yu Sebastian Cherng+*
cherng@nyu.edu

*Department of Sociology
University of Pennsylvania
3718 Locust Walk
Philadelphia, Pennsylvania 19104 USA

+*Department of Humanities and Social Sciences
New York University
Steinhardt School of Culture, Education, and Human Development
82 Washington Square East,
New York, New York 10003 USA

Introduction

*The authors' participation in the China Household Ethnic Survey is supported by grants from the University of Pennsylvania's School of Arts and Sciences Opportunity Fund and the University of Pennsylvania Provost's Global Engagement Fund.

Facility with a country's dominant language can be viewed as a key form of "linguistic capital", and has a role to play in processes of social stratification and mobility. This paper begins to explore the distribution of linguistic capital and the implications of possessing linguistic capital for economic opportunity among a historically disadvantaged group of young adults. After a discussion of the ways that linguistic capital might shape economic opportunity, we investigate the associations of language facility with information access and economic outcomes among rural young adults in western minority areas of China.

Specifically, we address three questions: 1) What is the reported degree of facility in standard Mandarin and minority languages for young adults in western minority areas? 2) Is information access, as reflected by Internet use, associated with reported language facility in standard Mandarin and minority languages? and, 3) Are young adult economic outcomes associated with language facility in standard Mandarin and minority languages?

Motivation

Facility with a country's dominant language may be viewed as a form of linguistic capital--a social resource that is valuable in gaining access to a nation's goods and positions (Smits and Gündüz-Hoşgör 2003; see Bourdieu 1992). Forms of capital—linguistic, cultural, social, economic, or symbolic—can, to some degree, be converted to other forms (Bourdieu 1986). That is, linguistic capital in the form of perfect mastery of a standard official language or dialect could be "transferred into other forms of capital like economic or social capital and thus help the 'legitimate speakers of the legitimate language' on their path to social success." (Smits and Gündüz-Hoşgör 2003, 830).

Language can create "symbolic barriers" to accessing social and economic resources if imperfect use of the dominant language or use of non-standard dialects triggers negative perceptions or discrimination (Smits and Gündüz-Hoşgör 2003). For example, in the United States, in a longitudinal study of 444 adolescent Chinese Americans, self-reports of speaking with an accent were associated with perceptions of being a perpetual foreigner and other forms of chronic daily discrimination, as well as increased risk of depressive symptoms (Kim et al. 2011). Also in the United States, in an audit study of housing market inquiries among speakers of White Middle-Class English, Black Accented English, and Black English Vernacular, Massey and Lundy (2001) found evidence of racial discrimination based on language. The authors found that, compared to speakers of White Middle-Class English, speakers of Black Accented and Vernacular English were less likely to speak with a rental agent, less likely to be told that a particular unit was available, and more likely to pay application fees and have credit worthiness mentioned as a potential problem in lease qualification. Accents have also been linked to labor market outcomes (Kalin and Rayko 1978; Kalin, Rayko, and Love 1980). For example, one study using data from the 1979 National Chicano Survey in the United States found that Mexican Americans who spoke English with an accent, independent of proficiency, earned significantly lower wages than their peers who spoke non-accented English (Alberto, Bohara, and Saenz 1993).

For those who do not speak the national language at all, or who experience extremely limited facility, there may be other "direct barriers" to accessing public resources, as argued by Smits and Gündüz-Hoşgör (2003, 830):

> *People who are not able to speak a country's dominant language have less access to written and spoken sources, cannot fulfill official jobs, are restricted in their relationships to their own social group, and depend on others for information that*

may be important for them. The negative consequences of their lack of linguistic capital can be expected to be stronger than for individuals who may speak a dialect of the dominant language, but to whom the dominant language is at least understandable.

In Turkey, Smits and Gündüz-Hoşgör (2003) found non-Turkish speaking women to be less employed in the formal economy, to have husbands with lower educational levels, and to have lower family incomes. In the United States, a body of literature has documented that immigrants who have greater proficiency in English enjoy both better access to jobs (lower unemployment rates) and higher wages (Dávila, Mora, and González 2011; Neidert and Tienda 1984; Stolzenberg and Tienda 1997).

Among "direct barriers" to economic opportunity, access to information is likely to play a key role, and a key source of information lies in the ability to access the Internet. Internet access and utilization have tended to be highly stratified across socio-economic groups.[1] Language itself is a likely stratifier of Internet access. For example, Kralisch and Mandl (2006, 1,8,9) employ an analysis of data collected from a multilingual e-health website to argue that language may represent a double barrier to information access. Most obviously, the language in which the information is presented influences who accesses the website, but there are also fewer webhosts and links available for Internet users from language groups that are less well represented on the Internet.

[1] In the United States, for example, individuals who have higher incomes (DiMaggio et al. 2001; Hargittai 1999), have higher levels of education (Bucy 2000; Coley, Cradler, and Engel 1997; DiMaggio and Hargittai 2001), are male (Ono and Zavodny 2003; Wasserman and Richmond-Abbott 2005; Weiser 2000), and are White (Hoffman and Novak 1998; Hoffman, Novak, and Schlosser 2001) have more access to, as well as fluency with, the Internet.

In this paper, we treat information access and economic outcomes in young adulthood as important indicators of opportunities for economic advancement. We investigate the relationship between language facility and these outcomes for rural young adults from majority and minority backgrounds in western minority areas of China. While we are not able to establish causal relationships in this paper, we hope that a sketch of the associations among language, information access and economic outcomes can be a first step toward illuminating the links between language and economic opportunity for historically-disadvantaged and still-impoverished populations coming of age in a time of rapid development, rising inequality, and increasing digital connection.

Specifically, we address three questions: 1) What is the reported degree of facility in standard Mandarin and minority languages for young adults in western minority areas? 2) Is information access, as reflected by Internet use, associated with reported language facility in standard Mandarin and minority languages? 3) Are young adult economic outcomes associated with language facility in standard Mandarin and minority languages?

Data and methods

Data

Data for our analysis comes from the China Household Economic Survey (CHES). The survey asks household heads to answer questions on all household members, which include non-resident members.[2] To highlight ways that language may link to opportunity for a disadvantaged

[2] The National Bureau of Statistics, on other household surveys, also asks household heads to answer questions about non-resident members (非常住人).

population, we focus on rural young adults in western China, for whom the transition to work would have been relatively recent. We use a sample of rural young adults ages 22 to 31. From this pool, we also draw a smaller sample of individuals belonging to an official ethnic minority group for analyses examining bilingualism. Table 1 shows descriptive statistics for the variables used in our analysis.

<Table 1 about here.>

Measurement

We use three groups of outcome measures in our analyses: language facility, Internet usage, and employment and migration outcomes. Our analyses also include a series of demographic and location measures. All measures come from the rural household questionnaire.

Language and linguistic capital measures

STANDARD MANDARIN FACILITY SCALE

Our primary measure of linguistic capital is a *Standard Mandarin Facility* scale, which is a scale of two items: the respondent's ability to communicate verbally in Mandarin and to read and write in Mandarin. The scale is created by summing the values on the two items and then dividing by the number of items with non-missing values over which the sum is calculated. The standard Mandarin facility scale is constructed by using standardized values (the mean is 0 and variance 1) of the individual items. Higher values on this scale reflect greater standard Mandarin facility.

STANDARD MANDARIN FACILITY = BASIC+

Standard Mandarin facility = basic+ is a categorical variable representing whether the individual reports strong ability or the ability to "basically" speak standard Mandarin, in contrast with simple or no communication ability in standard Mandarin. To preserve the maximum number of observations possible in our analyses, we include a third category that captures individuals who are missing observations on spoken Mandarin facility.[3]

MINORITY LANGUAGE FACILITY SCALE

Another measure used is the *minority language facility scale*. The first item asked ethnic minority individuals who reported having an ethnic minority spoken language about their ability to communicate in their minority language.[4] The second item asked ethnic minority individuals who reported having an ethnic minority written language about their ability to read and write in their minority language.[5] The *minority language facility* scale is created in the same fashion as the *standard Mandarin facility* scale. This question was only asked of ethnic minority individuals, so we are not able to consider Han individuals' minority language facility.

MINORITY LANGUAGE FACILITY=BASIC+

---

[3] Uygurs comprise the majority of ethnic minorities who are missing on observations of spoken Mandarin facility.

[4] On the rural household survey, ethnic minorities (and not Han) were asked two questions about communication ability in their minority language. One question asked ethnic minority respondents whether they had their own minority language ("是否有本民族的语言? "). Respondents who answered yes were then asked to answer another question that which asked about their communication ability in their minority language. Results are presented here for all those who answered the second question.

[5] Similar to the skip pattern for communication ability in ethnic minority language, only ethnic minorities were asked a set of questions about their reading and writing ability in their minority language. One question asked ethnic minority respondents whether they had their own written minority language ("是否有本民族的文字?"). Respondents who answered yes were then asked about their reading and writing ability in their minority language.

Minority language facility = basic+ is a categorical variable representing whether ethnic minority individuals report being a strong speaker of their own minority language or being "basically able" to communicate in the language, in contrast to reporting simple or no communication ability in their ethnic minority language. A category is also included to identify individuals who are missing observations on spoken minority language facility.[6]

Information access measures

USE INTERNET SOMETIMES OR OFTEN

Use Internet sometimes or often is a single binary variable that is coded 0 if the respondent reports never or almost never accessing the Internet, and 1 if the respondent reports accessing the Internet sometimes or often.

Economic outcomes measures

The third category of analytic variables we consider is economic outcomes, which includes employment- and migration- related dimensions. These variables draw from survey questions that are asked of all household members who have worked in 2011.

ENGAGE IN LOCAL AGRICULTURAL EMPLOYMENT, ENGAGE IN LOCAL WAGE EMPLOYMENT

Engage in local agricultural employment and *engage in local wage employment* are coded 0 if the individual does not engage in the particular type of employment and 1 if the respondent does.

[6] The Hui comprise the majority of ethnic minorities who are missing on observations of spoken minority language facility.

PERSONAL INCOME

Log personal income is the log of the sum of an individual's personal income and transfer income, for those with non-zero personal income.[7]

MIGRATION

For those who report working, *migration* is coded 0 if the individual has no migration experiences and 1 if the individual has already migrated.

Background measures

In addition to the key analytic variables emphasized in this analysis, we utilize a set of variables measuring background characteristics: age, gender, ethnic group membership and minority status, province of residence, and education.

AGE

The respondent's *age* is measured in years.

FEMALE

Female is coded 0 if male, 1 if female.

ETHNIC GROUP

[7] Personal income includes income from local non-agricultural wage employment, net revenue of local non-agricultural self-employment, migrant earnings (wage, net revenue), household transfer income, income from employment by state/collective for public work, and other working income. Transfer income includes income from allowance for retired workers and the new rural pension program.

A series of indicator variables was included to capture membership in an officially-recognized *ethnic group* with a sample size greater than 100 individuals. Large ethnic groups included here are Han, Mongolian, Hui, Tibetan, Uygur, Miao, Zhuang, Dong, Yao, and Tujia. A final indicator variable represents the remaining ethnic groups in an Other category (including Bouyei, Korean, Manchu, Kazak, Dai, Li, Lisu, Wa, She, Dongxiang, Jingpo, Tu, Mulao, Salar, Pumi, Tajik, Uzbek, Russian, and Oroqen).

MINORITY STATUS

A binary measure is used to represent whether the respondent is a *member of a minority group* (coded 0 if Han and 1 if an ethnic minority).

PROVINCE

We use the term province as shorthand to include province-level administrative units, whether the units are officially designated as provinces or as autonomous regions. Provinces included in the sample are Hunan, Guizhou, and Qinghai, and Autonomous Regions included in the sample are Inner Mongolia, Guangxi, Ningxia, and Xinjiang. A series of dummy variables indicates *province* of residence.

EDUCATION

Education is a categorical measure that includes categories for no schooling, primary school, lower-middle school, vocational or technical upper-middle school, upper-middle school, and college or higher. An additional category is included to capture individuals who are missing observations on educational level.

Analytic approach

To address the reported degree of facility in standard Mandarin and minority languages for young adults, we employ generalized least squares models with random effects at the level of the household. We consider differences in analytic variables across ethnic groups and provinces, and incorporate controls for gender, age, and education. Results for standard Mandarin facility are reported for individuals of all ethnic groups. Results for minority language facility are reported only for members of ethnic minorities.

We investigate the relationship between language facility and Internet use with random effects logit models of Internet use that adjust for demographic and educational characteristics. In these analyses, we also consider whether standard Mandarin facility benefits minority and non-minority individuals differently. Finally, to investigate the association between language facility and economic outcomes, we estimate random effects regression and logit models of economic outcomes that adjust for demographic and educational controls.

Results

Facility in standard Mandarin and minority languages

<Figure 1 about here.>

Figure 1 shows the proportion of the analytic sample who report that they are "basically able" to speak Mandarin or better, by province and minority status among those ages 22 to 31. There are stark differences in both the level of reported standard Mandarin facility and the nature of the minority-majority gap across provinces. For example, in Xinjiang, virtually all of the sample who are not minorities report basic or strong standard Mandarin facility, while just 20 percent of those who are members of minority groups do so. These figures, respectively,

represent the highest and lowest facility across each province-ethnic status grouping, and the largest ethnic gap. In contrast, in some provinces in the CHES sample—Hunan, Guangxi, and Guizhou—minority respondents report similar or greater facility than non-minority respondents with standard Mandarin.[8] In Guizhou, Qinghai, and Ningxia, 20 percent or more of young adults of both minority and non-minority backgrounds reported low levels of Mandarin facility.

Low facility with standard Mandarin amongst non-minority rural youth in some provinces highlights the importance of what might be called the invisible minority languages of China: local topolects (方言, *fangyan*) considered to be within the Chinese language. Use of non-standard local topolects could be a source of economic disadvantage amongst rural youth. Disadvantage could arise in the form of "symbolic" barriers to economic opportunity, if use of spoken topolects signals stigmatized, disadvantaged rural origins or poor quality of education, given that standard Mandarin is usually the language of school instruction for non-minority youth. Disadvantage could arise in the form of "direct" barriers, if local topolects are mutually unintelligible with standard Mandarin, heightening the challenge of communication. Children without Mandarin facility at home may face extra challenges in obtaining human capital, as the critical early school transition may be particularly challenging for students starting to learn in a different topolect.

<Table 2 about here.>

[8] Using two-sample tests for proportions, the difference between the proportion of minorities and Han in Guizhou who report speaking basic+ Mandarin is statistically significant at the $p < 0.05$ level. The differences between these groups in Hunan and Guangxi are not statistically significant.

Table 2 shows regression coefficients from linear regression models estimating the standard Mandarin language facility scale for the full sample (Panel A), and standard Mandarin and minority language facility for all respondents who were members of ethnic minority groups (Panel B). Table 2 shows that being older and being female are associated with reporting less language facility in Mandarin, and more education is associated with reporting greater facility. These statements are true both for the full sample in Panel A and for the sample of members of ethnic minority groups in Panel B. In Panel B, for minority language facility, as for the case of standard Mandarin, education is positively associated, but unlike the case for standard Mandarin, gender and age are not associated.

Panel A also shows coefficients for individual ethnic groups. Compared to the Han, there are significant linguistic disadvantages in standard Mandarin across specifications for Tibetans and Uygurs, meaning that these disadvantages persist net of province and educational attainment. In the Tibetan case, the coefficient for Tibetan in Model 3, which adjusts for place and education, is reduced by about half compared to that in Model 1, suggesting that place and educational composition are part of the story.[9] For Mongolians, there is not a significant difference overall, but a significant disadvantage emerges net of location and persists net of education. For Hui, a disadvantage exists overall, and persists in the specification with geographic controls, but is insignificant in the specification with education included in the model. For the Miao and Dong, a disadvantage exists overall, but not with provincial and education controls. No significant

[9] Using the postestimation command for seemingly unrelated estimation (the SUEST and TEST commands in Stata) after running Models 1 and 3 from Table 2 using the REG command, we find that the coefficient for Tibetans is significantly different in Model 3 than in Model 1.

disadvantages in self-reported language facility are observed in these specifications for Zhuang, Yao, and Tujia.

Results show that non-trivial numbers of rural young adults in the sample, overall—majority and minority--report lack of facility with Mandarin. Members of some, but not all, minority groups experience relative disadvantages in facility with standard Mandarin, which may be associated with educational infrastructure, the climate for intercultural communication, geographic segregation within provinces, and other aspects of regions. These aspects likely include other dimensions of infrastructure and the degree of ethno-linguistic fractionalization in the region.

Mandarin, minority languages and information access

<Figure 2 about here.>

Figure 2 shows the proportion of individuals who report occasionally or often using the Internet, by province and minority status. In the sample, there is considerable variation in Internet use for residents of different provinces, overall. There is a generalized pattern in the sample in which members of minority groups in the sample are less likely to be occasional or frequent users of the Internet than their Han counterparts in the same province, but the gaps vary tremendously. For this sample of rural young adults, Xinjiang again stands out with the highest proportion of Internet users across all of the province-minority status groups among non-minorities, and the lowest proportion across all of the province-minority status groups among minorities.

<Table 3 about here.>

Does language play a role in information access? Table 3 presents regression coefficients from logistic regression models estimating use of the Internet. Focusing first on analyses with the full sample in Panel A, Model 1 includes age, gender, province, educational attainment, and minority group membership, and shows, overall, a pattern of less Internet use among older people relative to younger people, females relative to males, minorities relative to the majority, and the less educated relative to the more educated. The odds-ratio associated with minority status compared to Han status indicates about a 75 percent reduction in odds of using the Internet associated with minority status, net of other variables in the model.[10] Compared to residents of the reference province of Inner Mongolia, residents of Qinghai and Xinjiang report significantly less Internet use.

Model 2 adds to this baseline model the scale of facility with standard Mandarin, the coefficient for which is highly significant and positive. Adjusting for language facility in this specification, the odds-ratio associated with minority status is 0.42, indicating about a 58 percent reduction in odds of accessing the Internet, relative to non-minorities.[11] Model 3 adds an interaction term between minority status and language to address whether language facility might matter differently for members of minority groups, but the term is insignificant. Finally, Model 4 (in Panel B) focuses on members of minority groups, and considers language facilities in Mandarin and in the minority language. An interaction term tests whether strength in both minority language and Mandarin might be of extra benefit for Internet access. Model 4 shows a significant effect only for Mandarin facility. In summary, Mandarin facility, along with

[10] e^-1.40 = 0.25.
[11] e^-0.87 = 0.42.

education, is a significant correlate of Internet use, for both majority and minority members. Minority language facility is not associated with Internet use.

Language and economic outcomes

Table 4 shows coefficients from models of young adult economic outcomes: local agricultural employment, local wage employment, personal income and migration experience. Panel A focuses on employment-related outcomes and Panel B, on migration outcomes. Across the specifications presented here, age is associated with a higher likelihood of working locally and lower likelihood of reporting migration experience. Women are more likely than men to engage in local agriculture, less likely than men to be engaged in local wage employment, have lower incomes than men, and are less likely to report migration experience than men. Net of factors included in the models, minorities are more likely to engage in agriculture, have lower personal income, and may be less likely to migrate, though this latter result is only marginally significant and the significance dissipates after controlling for language facility. Residents of all provinces save Xinjiang are less likely than those in the reference province of Inner Mongolia to report agricultural employment, and there are geographic differences in personal income and migration experience, as well. Educational attainment is generally associated with lower likelihood of agricultural employment, greater likelihood of wage employment, more personal income, and, in specification 4a, migration experience.

Adjusting for other variables in the models, language facility in standard Mandarin is associated with significantly reduced likelihood of local agricultural employment. Standard Mandarin facility is not associated with local wage employment, but is positively associated with personal income. Panel B shows logit models distinguishing those who report migration experience from those who report no migration. Model 4a shows a specification without

language facility in standard Mandarin, and Model 4b shows a specification with language facility. Two key findings emerge. Net of other variables in the model, language facility in Mandarin is associated with migration experience. Moreover, education shows a significant, positive association with migration in Model 4a, but is no longer a significant, positive predictor of migration experience once language facility is introduced into the model in Model 4b. This pattern suggests that the observed education effect on migration, which is an important potential avenue of economic mobility for impoverished rural youth, may operate in part through the linguistic capital acquired, or not acquired, in school.

<Tables 4 and 5 about here.>

Table 5 shows estimates of language facility in Mandarin and minority languages for the minority subsample, with age, gender, province, and education as controls. Table 5 shows the same set of outcomes and predictors displayed in Table 4. Indicators for "basically speak" or better Mandarin ability and "basically speak" or better minority language facility are included. Results show that for minority populations, Mandarin facility, but not minority language facility, is associated with economic rewards. Even after adjusting for education, which is associated with language facility, those who report Mandarin facility are less likely to be employed in local agriculture, earn higher incomes, and are more likely to report migration experience, compared to their less-fluent counterparts. In contrast, minority language facility shows no significant positive associations with economic outcomes, and, in fact, facility with a minority language is actually associated with a *penalty* in personal income.

Conclusions

Facility with a country's dominant language, a key form of linguistic capital, has a role to play in processes of social stratification and mobility, and this role is poorly understood. We have sought, in this paper, to explore access to this form of linguistic capital, and the implications of possessing linguistic capital, for a group of young adults who have been historically disadvantaged: rural young adults in western minority areas. Three main results emerge. First, there is a great deal of variability in linguistic capital, defined as standard Mandarin facility, across provinces and ethnic groups covered in the CHES sample. The greatest gap appears in Xinjiang, where Han residents have very high facility in standard Mandarin, and where minority residents report very low facility. In some provinces in the CHES sample, there are minimal differences between majority and minority populations. Second, standard Mandarin facility is tied to information access, in the form of Internet use. Facility in minority languages is not. Third, Mandarin facility, but not minority language facility, is linked to economic opportunity in young adulthood.

While the complex theoretical interrelationships among these variables preclude causal statements about the impact of linguistic capital on information access and employment outcomes, the patterns revealed suggest some insights about language and inequality. First, Mandarin facility and minority language facility are associated with significantly different employment patterns in young adulthood. In Bourdieu's conceptualization, "all speech is produced for and through the market to which it owes its existence and its most specific properties." (Bourdieu 1992, 76; see discussion of this quote in Loos 2000, 38) There are different "markets" that endow particular linguistic products with different valuations, and "each speaker in a linguistic community possesses a certain quantity of linguistic capital which allow[s] him to produce expressions which are highly valued on a particular market" (Loos 2000, 38).

Findings suggest that those with high levels of Mandarin facility are operating in distinct linguistic and economic markets from those with high levels of facility in minority languages. These distinct patterns in early adulthood are likely to cumulate and solidify across the life course.

Second, Mandarin facility is associated with education, such that those without human capital are at greater likelihood of lacking linguistic capital. Moreover, those most vulnerable to poor economic outcomes—those who are poorly educated and without linguistic capital—are also more likely to lack access to information. The cohesion of multiple sources of disadvantage in the lives of these young adults represents the flip side of the convertability of different forms of capital—linguistic, cultural, human, economic, and symbolic. Drawing on Bourdieu to write about the European case, Loos (2000, 39) gives the example, "a good education (human and cultural capital) means speaking a foreign language (linguistic capital), [which] may help getting a good job at a multinational which pays well (economic capital) and gaining prestige (symbolic capital)." In the current case, for rural young adults, a "good education" and its implied human and cultural capital may be signaled in part by the capacity to operate in standard Mandarin. For both speakers of local topolects and minority languages, the lack of standard Mandarin may engender doubts in the labor market about the quality of candidates' other forms of capital—human and cultural—and thus harm their "market" value.

The measures of language facility that can be gleaned from omnibus household surveys are necessarily crude, and this problem is a significant limitation in the current paper. Yet, the pattern of findings suggests the common-sense insight that linguistic capital is likely to be importantly linked to stratification processes in China, as has been documented elsewhere. Further work is needed to illuminate the role of local contexts—such as educational

infrastructure, the climate for inter-cultural communication, ethno-linguistic composition, and the nature of local labor markets—in shaping relationships among language, information, and economic opportunity. In addition, detailed work on both the production of linguistic capital and nature of labor market disparities associated with language facility would add fruitfully to our understanding of the potential role of language in stratification processes in China. Finally, non-minority young adults in some cases also experience low levels of linguistic capital. This finding suggests the possibility that speakers of local topolects may experience symbolic or direct barriers to economic opportunity in similar ways to speakers of official minority languages, but these "invisible languages" are difficult to document with current data. Further research could fruitfully explore topolects as sources of advantage and disadvantage in education and labor markets.

References cited

Alberto, Davila, Alok K. Bohara, and Rogelio Saenz. 1993. "Accent Penalties and The Earnings of Mexican Americans."." *Social Science Quarterly* 74 (4): 902–916.

Bialystok, Ellen, and Raluca Barac. 2012. "Emerging Bilingualism: Dissociating Advantages for Metalinguistic Awareness and Executive Control." *Cognition* 122 (1): 67–73.

Bourdieu, Pierre. 1986. "The Forms of Capital." In , edited by John G. Richardson, 241–258. Handbook of Theory and Research for the Sociology of Education. New York: Greenwood Press.

———. 1992. *Language and Symbolic Power*. Cambridge: Polity Press.

Bucy, Erik P. 2000. "Social Access to the Internet." *The Harvard International Journal of Press/Politics* 5 (1): 50–61.

Chitiri, Helena-Fivi, and Dale M. Willows. 1997. "Bilingual Word Recognition in English and Greek." *Applied Psycholinguistics* 18 (02): 139–156.

Coley, R. J., John Cradler, and Penelope K. Engel. 1997. "Computers and Classrooms: The Status of Technology in US." *Schools. Princeton, NJ: Educational Testing Service*. http://integratingthearts.yolasite.com/resources/ED412893.pdf.

Dávila, Alberto, Marie T. Mora, and Rebecca González. 2011. "English-Language Proficiency and Occupational Risk Among Hispanic Immigrant Men in the United States." *Industrial Relations: A Journal of Economy and Society* 50 (2): 263–296.

DiMaggio, Paul, and Eszter Hargittai. 2001. "From the 'digital Divide'to 'digital Inequality': Studying Internet Use as Penetration Increases." *Princeton University Center for Arts and Cultural Policy Studies, Working Paper Series Number* 15. http://www.maximise-ict.co.uk/WP15_DiMaggioHargittai.pdf.

DiMaggio, Paul, Eszter Hargittai, W. Russell Neuman, and John P. Robinson. 2001. "Social Implications of the Internet." *Annual Review of Sociology* 27 (January 1): 307–336.

Garcia, Georgia Earnest. 1998. "Mexican-American Bilingual Students' Metacognitive Reading Strategies: What's Transferred, Unique, Problematic?." In *National Reading Conference Yearbook*, 47:253–63. http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ587439.

Hakuta, Kenji. 2011. "Educating Language Minority Students and Affirming Their Equal Rights Research and Practical Perspectives." *Educational Researcher* 40 (4): 163–174.

Hargittai, Eszter. 1999. "Weaving the Western Web: Explaining Differences in Internet Connectivity among OECD Countries." *Telecommunications Policy* 23 (10): 701–718.

Hoffman, Donna L., and Thomas P. Novak. 1998. "Bridging the Digital Divide: The Impact of Race on Computer Access and Internet Use." http://www.ehcweb.ehc.edu/faculty/ljcumbo/PDFS/hoffman_novak.pdf.

Hoffman, Donna L., Thomas P. Novak, and Ann E. Schlosser. 2001. "The Evolution of the Digital Divide: Examining the Relationship of Race to Internet Access and Usage over Time." *The Digital Divide: Facing a Crisis or Creating a Myth*: 47–97.

Kalin, Rudolf, and Donald S. Rayko. 1978. "Discrimination in Evaluative Judgments against Foreign-Accented Job Candidates." *Psychological Reports* 43 (3f): 1203–1209.

Kalin, Rudolf, Donald S. Rayko, and Norah Love. 1980. "The Perception and Evaluation of Job Candidates with Four Different Ethnic Accents." *Language: Social Psychological Perspectives*: 197–202.

Kim, Su Yeong, Yijie Wang, Shiying Deng, Rocio Alvarez, and Jing Li. 2011. "Accent, Perpetual Foreigner Stereotype, and Perceived Discrimination as Indirect Links between

English Proficiency and Depressive Symptoms in Chinese American Adolescents." *Developmental Psychology* 47 (1): 289–301. doi:10.1037/a0020712.

Kralisch, Anett, and Thomas Mandl. 2006. "Barriers to Information Access across Languages on the Internet: Network and Language Effects." In *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on*, 3:54b–54b.

Loos, E. 2000. "Language Choice, Linguistic Capital and Symbolic Domination in the European Union." *Language Problems & Language Planning* 24 (1) (December 1): 37–53. doi:doi:10.1075/lplp.24.1.04loo.

Massey, Douglas S., and Garvey Lundy. 2001. "Use of Black English and Racial Discrimination in Urban Housing Markets New Methods and Findings." *Urban Affairs Review* 36 (4) (March 1): 452–469. doi:10.1177/10780870122184957.

Nagy, William, Virginia W. Berninger, and Robert D. Abbott. 2006. "Contributions of Morphology beyond Phonology to Literacy Outcomes of Upper Elementary and Middle-School Students." *Journal of Educational Psychology* 98 (1): 134.

Neidert, Lisa J., and Marta Tienda. 1984. "Converting Education into Earnings: The Patterns among Hispanic Origin Men." *Social Science Research* 13 (4): 303–320.

Ono, Hiroshi, and Madeline Zavodny. 2003. "Gender and the Internet*." *Social Science Quarterly* 84 (1): 111–121.

Smits, Jeroen, and Ayşe Gündüz-Hoşgör. 2003. "Linguistic Capital: Language as a Socio-Economic Resource among Kurdish and Arabic Women in Turkey." *Ethnic and Racial Studies* 26 (5) (September 1): 829–853. doi:10.1080/0141987032000109050.

Stolzenberg, Ross M., and Marta Tienda. 1997. “English Proficiency, Education, and the Conditional Economic Assimilation of Hispanic and Asian Origin Men.” *Social Science Research* 26 (1): 25–51.

Wasserman, Ira M., and Marie Richmond-Abbott. 2005. “Gender and the Internet: Causes of Variation in Access, Level, and Scope of Use*.” *Social Science Quarterly* 86 (1): 252–270.

Weiser, Eric B. 2000. “Gender Differences in Internet Use Patterns and Internet Application Preferences: A Two-Sample Comparison.” *CyberPsychology and Behavior* 3 (2): 167–178.

Table 1. Descriptive Statistics for Variables Used in Analyses

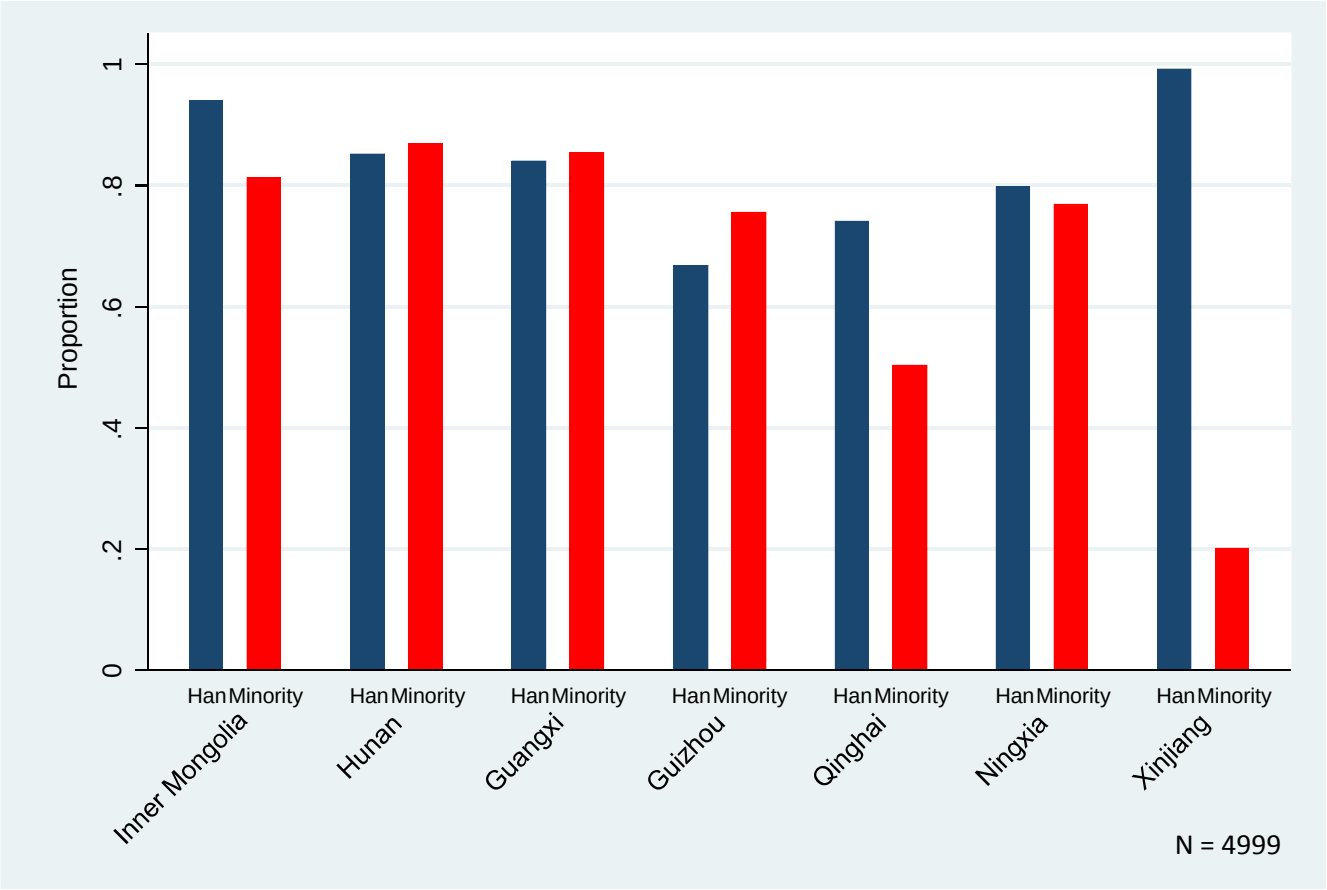
	Sample=Han			Sample=Minority			Sample=Total		
	Mean / Proportion	SD	N	Mean / Proportion	SD	N	Mean / Proportion	SD	N
Language and linguistic capital measures									
Language facility scales (speaking, reading, and writing)									
Standard Mandarin facility scale	0.60	0.66	1747	0.24	0.82	3284	0.37	0.79	5153
Minority language facility scale	--			0.12	0.89	3240	--		
Standard Mandarin speaking facility=basic+			1792			3624			5416
Below basic spoken Mandarin	0.16		289	0.26		930	0.23		1219
Basic+ spoken Mandarin	0.80		1441	0.65		2339	0.70		3780
Missing	0.03		62	0.10		355	0.08		417
Minority language speaking facility=basic+ in …						3624			
Below basic spoken minority language	--			0.22		799	--		
Basic+ spoken minority language	--			0.64		2330	--		
Missing	--			0.14		495	--		
Information access measures									
Internet usage									
Sometimes or often (ref: never, almost never)	0.50		1718	0.36		3222	0.41		4940
Economic outcomes measures									
Local agricultural employment	0.52		1792	0.57		3624	0.55		5416
Local wage employment	0.13		1792	0.13		3617	0.13		5409
Logged personal income	9.27	1.08	1102	9.10	1.17	2257	9.16	1.14	3359
Migration (ref: have not migrated)	0.36		1788	0.36		3606	0.36		5394
Demographic and location measures									
Age	25.98	2.92	1792	26.06	2.87	3624	26.03	2.89	5416
Female	0.47		1792	0.46		3624	0.46		5416
Ethnic group			1792			3624			5416
Han				--			33.09		1,792
Mongolian	--			0.05		163	0.03		163
Hui	--			0.13		485	0.09		485
Tibetan	--			0.07		252	0.05		252
Uygur	--			0.12		440	0.08		440
Miao	--			0.23		818	0.15		818
Zhuang	--			0.11		383	0.07		383
Dong	--			0.13		479	0.09		479
Yao	--			0.04		140	0.03		140

Table 1. Descriptive Statistics for Variables Used in Analyses

	Sample=Han			Sample=Minority			Sample=Total		
	Mean / Proportion	SD	N	Mean / Proportion	SD	N	Mean / Proportion	SD	N
Tujia	--			0.04		144	0.03		144
Other	--			0.09		320	0.06		320
Member of a minority group	--			--			0.67		5416
Province			1792			3624			5416
Inner Mongolia	0.18		326	0.05		168	0.09		494
Hunan	0.12		218	0.18		655	0.16		873
Guangxi	0.21		374	0.19		685	0.20		1,059
Guizhou	0.10		172	0.21		749	0.17		921
Qinghai	0.15		276	0.15		532	0.15		808
Ningxia	0.16		287	0.09		327	0.11		614
Xinjiang	0.08		139	0.14		508	0.12		647
Education			1792			3624			5416
No schooling	0.02		32	0.05		197	0.04		229
Primary school	0.10		179	0.20		733	0.17		912
Lower-middle school	0.54		970	0.54		1,940	0.54		2,910
Vocational or technical upper-middle school	0.06		100	0.04		134	0.04		234
Upper-middle school	0.11		199	0.08		284	0.09		483
College or higher	0.14		255	0.07		269	0.10		524
Missing	0.03		57	0.02		67	0.02		124

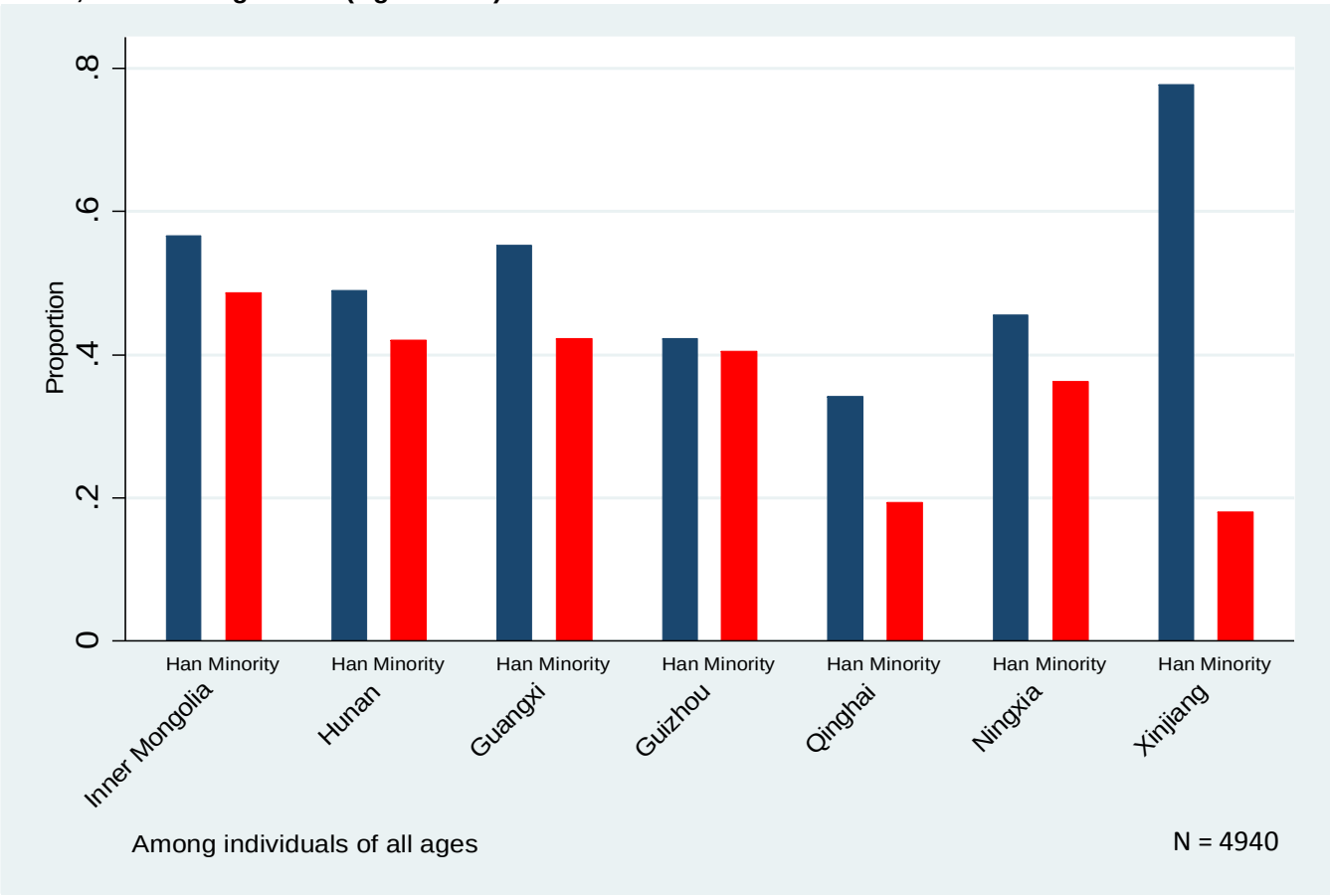
Source: China Household Economic Survey

Figure 1. Proportion Who Report "Basically" Speaking Mandarin or Better (BASIC+), by Province and Minority Status, Rural Young Adults (Ages 22-31)



Source: China Household Economic Survey

Figure 2. Proportion Who Report Occasionally or Often Using the Internet, by Province and Minority Status, Rural Young Adults (Ages 22-31)



Source: China Household Economic Survey

Table 2. Coefficients from Regression Models Estimating Language Facility, Rural Young Adults (Ages 22-31)

	Panel A: Sample=All			Panel B: Sample=Official Minorities	
	Standard Mandarin			Standard Mandarin	Minority Language
	(1)	(2)	(3)	(4)	(5)
Age	-0.03***	-0.03***	-0.01***	-0.01**	-0.00
Female	-0.15***	-0.15***	-0.07***	-0.10***	-0.01
Ethnic Group (reference: Han)					
Mongolian	-0.08	-0.41***	-0.42***		
Hui	-0.33***	-0.23***	0.01		
Tibetan	-0.81***	-0.55***	-0.38***		
Uygur	-1.63***	-1.76***	-1.53***		
Miao	-0.17***	-0.07*	-0.01		
Zhuang	-0.03	0.04	0.02		
Dong	-0.12***	-0.01	0.01		
Yao	-0.06	-0.02	0.03		
Tujia	-0.07	-0.07	-0.01		
Other	-0.41***	-0.31***	-0.14***		
Province (ref: Inner Mongolia)					
Hunan		-0.37***	-0.33***	0.09	-0.94***
Guangxi		-0.46***	-0.38***	0.10	-0.26***
Guizhou		-0.57***	-0.48***	-0.04	-0.52***
Qinghai		-0.71***	-0.54***	-0.33***	-0.25***
Ningxia		-0.36***	-0.32***	0.18**	-1.38***
Xinjiang		-0.25***	-0.36***	-1.27***	0.65***
Education (reference: no schooling)					
Primary school			0.61***	0.61***	0.18***
Lower-middle school			0.95***	0.99***	0.21***
Vocational or technical upper-middle school			1.19***	1.25***	0.28***
Upper-middle school			1.13***	1.15***	0.25***
College or higher			1.49***	1.53***	0.34***
Missing			0.91***	0.92***	-0.11
Constant	1.48***	1.87***	0.23***	-0.20*	0.31***
Observations	5024	5024	5024	3278	3234

*** p<0.01, ** p<0.05, * p<0.1

Source: China Household Economic Survey

Table 3. Coefficients from Logistic Regression Models Estimating Use of Internet, Rural Young Adults (Ages 22-31)

	Panel A: Sample=All			Panel B: Sample=Official Minorities
	(1)	(2)	(3)	(4)
Age	-0.41***	-0.44***	-0.44***	-0.48***
Female	-1.57***	-1.50***	-1.49***	-1.69***
Province (ref: Inner Mongolia)				
Hunan	-0.20	0.45	0.32	-0.69
Guangxi	0.35	1.28**	1.12*	-0.13
Guizhou	-0.03	1.23*	1.09*	0.28
Qinghai	-2.77***	-1.20*	-1.30**	-2.22**
Ningxia	-0.93	-0.20	-0.34	-1.91
Xinjiang	-1.78***	1.84**	1.88**	-2.22*
Education (reference: no schooling)				
Primary school	2.14***	0.91	0.89	1.62*
Lower-middle school	4.31***	2.26***	2.23***	3.27***
Vocational or technical upper-middle school	8.23***	5.59***	5.53***	7.82***
Upper-middle school	5.83***	3.21***	3.19***	5.21***
College or higher	11.70***	7.62***	7.58***	11.36***
Missing	4.63***	2.31*	2.26*	4.17**
Language facility				
Minority language facility scale				-0.93
Standard Mandarin facility scale		4.16***	3.72***	4.05***
Minority language facility scale * Standard Mandarin facility scale				0.40
Member of a minority group	-1.40***	-0.87***	-1.18***	
Member of a minority group * Standard Mandarin facility scale			0.61	
Constant	8.22***	7.53***	7.92***	6.61***
Observations	4933	4922	4922	2835

*** p<0.01, ** p<0.05, * p<0.1

Source: China Household Economic Survey

Table 4. Coefficients from Linear, Logistic, and Multinomial Logistic Regression Models Estimating Employment and Migration Outcomes, Rural Young Adults (Ages 22-31)

	Panel A: Logit or linear regression models of selected employment outcomes			Panel B: Logit Models of Migration (ref: has not migrated)	
	Engage in local agriculture employment	Engage in local wage employment	Log personal Income (OLS)		
	(1)	(2)	(3)	(4a)	(4b)
Age	0.12***	0.13***	0.00	-0.05***	-0.05***
Female	0.63***	-1.06***	-0.38***	-0.55***	-0.54***
Member of a minority group	0.35**	0.09	-0.11**	-0.12*	-0.10
Province (ref: Inner Mongolia)					
Hunan	-3.22***	0.12	0.25**	1.72***	1.72***
Guangxi	-3.02***	-0.27	0.11	1.93***	1.94***
Guizhou	-3.09***	-0.36	0.37***	1.84***	1.92***
Qinghai	-1.21***	0.09	0.74***	0.86***	1.01***
Ningxia	-1.65***	0.49	0.60***	1.29***	1.37***
Xinjiang	-0.54	0.45	0.20*	-0.30	0.06
Education (reference: no schooling)					
Primary school	0.26	0.77*	0.20*	0.13	0.01
Lower-middle school	-0.44	0.97**	0.38***	0.43**	0.17
Vocational or technical upper-middle school	-1.96***	1.54***	0.59***	0.68***	0.37
Upper-middle school	-0.35	1.27***	0.42***	0.48**	0.19
College or higher	-4.19***	0.73	0.79***	0.31	-0.08
Missing	-1.30**	1.51**	0.11	-0.11	-0.23
Standard Mandarin speaking facility=basic+	-1.29***	0.16	0.21***		0.86***
Missing standard Mandarin speaking facility=basic+	-0.41	0.19	-0.21**		0.24
Constant	-0.01	-6.51***	8.80***	-0.17	-0.72*
Observations	5416	5403	3359	5387	5387

*** p<0.01, ** p<0.05, * p<0.1

Source: China Household Economic Survey

Table 5. Coefficients from Linear, Logistic, and Multinomial Logistic Regression Models Estimating Employment and Migration Outcomes Among Ethnic Minorities, Rural Young Adults (Ages 22-31)

	Panel A: Logit or linear regression models of selected employment outcomes			Panel B: Logit Models of Migration (ref: has not migrated)
	Engage in local agriculture employment	Log personal Income (OLS)	Log personal Income (OLS)	
	(1)	(2)	(3)	(4)
Age	0.07***	0.13***	0.00	-0.04***
Female	0.79***	-1.03***	-0.38***	-0.63***
Province (ref: Inner Mongolia)				
Hunan	-3.08***	0.26	0.07	1.59***
Guangxi	-2.83***	-0.34	-0.14	1.99***
Guizhou	-3.03***	-0.39	0.17	1.90***
Qinghai	-0.67	0.22	0.48***	0.86***
Ningxia	-0.32	0.96*	0.16	0.72**
Xinjiang	0.38	0.80	0.01	-0.81**
Education (reference: no schooling)				
Primary school	0.20	0.44	0.19	-0.09
Lower-middle school	-0.55	0.65	0.39***	0.12
Vocational or technical upper-middle school	-1.65***	1.21**	0.58***	0.34
Upper-middle school	-0.48	0.65	0.38***	0.05
College or higher	-4.43***	0.81	0.73***	-0.40*
Missing	-1.46*	1.79**	0.02	-0.25
Minority language speaking facility=basic+	-0.15	-0.36	-0.16**	-0.01
Standard Mandarin speaking facility=basic+	-1.27***	0.19	0.22***	0.84***
Constant	1.34	-6.21***	9.04***	-0.77
Observations	3624	3611	2257	3600

*** p<0.01, ** p<0.05, * p<0.1

Source: China Household Economic Survey